

University of Groningen

Action Categorisation in Multimodal Instructions

van der Sluis, Ielka; Vergeer, Renate; Redeker, Gisela

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Final author's version (accepted by publisher, after peer review)

Publication date:

2018

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

van der Sluis, I., Vergeer, R., & Redeker, G. (2018). *Action Categorisation in Multimodal Instructions*. 31-36. Poster session presented at AREA - Annotation, Recognition and Evaluation of Actions, Miyazaki, Japan.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Action Categorisation in Multimodal Instructions

ABSTRACT

We present an explorative study for the (semi-)automatic categorisation of actions in Dutch multimodal first aid instructions, where the actions needed to successfully execute the procedure in question are presented verbally and in pictures. We start with the categorisation of verbalised actions and expect that this will later facilitate the identification of those actions in the pictures, which is known to be hard. Comparisons of and user-based experimentation with the verbal and visual representations will allow us to determine the effectiveness of picture-text combinations and will eventually support the automatic generation of multimodal documents. We used Natural Language Processing tools to identify and categorise 2,388 verbs in a corpus of 78 multimodal instructions (MIs). We show that the main action structure of an instruction can be retrieved through verb identification using the Alpino parser followed by a manual selection operation. The selected main action verbs were subsequently generalised and categorised with the use of Cornetto, a lexical resource that combines a Dutch Wordnet and a Dutch Reference Lexicon. Results show that these tools are useful but also have limitations which make human intervention essential to guide an accurate categorisation of actions in multimodal instructions.

Keywords: instructions, actions, verbs, categorisation, task structure.

RESEARCH QUESTION

How can the constituent actions in Dutch first aid procedures be identified and categorised by (semi-)automatic natural language analysis, so that the resulting action categories can be used to identify picture-text relations in multimodal instructions?

CORPUS

78 multimodal instructions (MIs) from the annotated PAT corpus¹

Published in two editions of *Het Oranje Kruis Boekje* (2011², 2016³)

Overlap in terms of the first aid tasks:

- 25 tasks appear in both editions (yielding 50 MIs),
- and 28 appear only in one edition.



Het Oranje Kruis

Het Oranje Kruis⁴ is a Dutch organisation that provides learning materials for first aid certification trainings.

¹ <http://www.rug.nl/let/pat>

² Het Oranje Kruis (2011). *Het Oranje Kruis Boekje*, De officiële handleiding voor eerste hulp. Thieme Meulenhoff, Amersfoort. ISBN 9789006921717.

³ Het Oranje Kruis (2016). *Het Oranje Kruis Boekje*, De officiële handleiding voor eerste hulp. Thieme Meulenhoff, Amersfoort. ISBN13 9789006410341.

⁴ <http://www.hetoranjekruis.nl/>

EXAMPLES

How to place a victim in the recovery position.



Figure 1: MI911 Placing a victim in the recovery position (Het Oranje Kruis Boekje, 2011).



Figure 2: Part of MI959 Placing a victim in the recovery position (Het Oranje Kruis Boekje, 2016).

METHOD

1. Selection of the verbs from MIs using Alpino (Van Noord, 2006)
2. Manual exclusion of modalised actions, negated actions, conditional actions, and warnings (see Table 1)
3. Selection of hyperonyms for the remaining verbs using Cornetto (Vossen et al., 2013)
4. Abstraction from hyperonyms using Cornetto 2.0 synset XML and other sources (see Table 2)
5. Further analysis using the Cornetto Demo of 21 hyperonyms with a frequency > 1% excluding semantically weak verbs (e.g., to be, to go, to have, to come)

BIBLIOGRAPHY

Van Noord, G. J. M. (2006). At Last Parsing Is Now Operational. In P. Mertens, C. Faron, A. Dister, & P. Watrin (Eds.), *TALN06. Verbum Ex Machina. Actes de la 13e conference sur le traitement automatique des langues naturelles*. (pp. 20–42). Leuven University Press.

Van der Sluis, I., Eppinga, A. & Redeker, G. (2017). Text-Picture Relations in Multimodal Instructions. In Asher, N. , Hunter, J. & Lascarides, A. (Eds.), *Proceedings of the workshop on Foundations of Situated or Multimodal Communication (FMSC) at IWCS 2017*.

Vossen, P., Maks, I., Segers, R., Van Der Vliet, H., Moens, M., Hofmann, K., Tjong Kim Sang, E. & De Rijke, M. (2013). Cornetto: a combinatorial lexical semantic database for Dutch. In *Essential Speech and Language Technology for Dutch*, (pp.165–184). Springer.

RESULTS

Table 1: Features and words used to identify modal verbs, negated actions, conditional actions and warnings in the parsed MI texts and the number and percentages of excluded verbs.

FEATURE	WORDS	EXCLUDED VERBS
Modal Verbs	kunnen (can, N=161), moeten (should, N=30), mogen (may, N=10)	221 (9.3%)
Explicit Negation	niet (not)	297 (12.4%)
Integrated Negation	geen (none, N=70), niemand (no one, N=14), nooit (never, N=7), niets (nothing, N=5)	96 (4.0%)
Other negation elements	alleen (only, N=49), maar (but, N=30), zonder (without, N=28), minder (less, N=25), enkel (just, N=10), hoogstens (at most, N=4), slechts (only, N=9), nauwelijks (barely, N=7), pas (only just, N=4), weinig (few, N=4), zelden (rare, N=2), moeilijk (hard, N=2)	174 (7.3%)
Conditionals	als (if, N=226), wanneer (when, N=170), zolang (as long as, N=8), indien (if, N=8)	408 (17.1%)
Warnings	voorkomen (to prevent, N=56), opletten (to pay attention, N=46), uitkijken (to watch out, N=4)	106 (4.4%)
Excluded verbs		977 (40.9%)
Main verbs		1,411 (59.1%)
Total		2,388 (100%)

Table 2: Hyperonym sources.

SOURCE	UNIQUE VERBS	TOTAL NUMBER OF VERBS
Cornetto 2.0 DB	127 (45.0%)	938 (66.5%)
Cornetto Demo	118 (41.8%)	365 (25.3%)
Van Dale Dictionary	25 (8.9%)	83 (5.9%)
Prefix stripping	12 (4.3%)	25 (1.7%)
Total	282 (100%)	1,411 (100%)

Table 3: Eight hyperonyms categories with their synsets, frequencies and corpus examples translated from Dutch to English.

	SYNSET CATEGORIES	FREQ.	TRANSLATED EXAMPLES (MI NUMBER)
1.	handelen (to do)	155 (11%)	Repeat these last steps until you are out of the danger zone. (903)
2.	veranderen (to change)	57 (4.0%)	Someone has immediately become seriously ill. (987) You take turns in resuscitating every 2 minutes.(958)
2.1	vastmaken (to attach)	61 (4.3%)	Attach the bandage with adhesive plaster or a bandage clip. (933)
2.2	bewerken (to manipulate)	28 (2.0%)	Prepare the breathing mask for use. (959)
2.2.1	dekken (to cover)	17 (1.2%)	Then you cover the wound with a sterile bandage. (971)
2.3	draaien (to turn)	31 (2.2%)	Place your hand on his forehead and tilt his head backwards. (911)
2.4	brengen (to bring)	23 (1.6%)	Put a stifled victim in a half-sitting position and support him. (980)
3.	houden (to hold, to keep)	32 (2.5%)	Hold his head in the position in which you found it.(979) Keep clinging clothing wet. (989)
4.	geven (to give)	26 (1.8%)	This is how you give enough support without squeezing. (973)
5.	voortbewegen (to propel)	20 (1.6%)	Slide both your arms under the victim’s armpits. (951)
5.1	verplaatsen (to move)	35 (2.5%)	Lift him by stretching your legs. (902)
5.1.1	plaatsen (to put, to place)	99 (7.0%)	Place the CPR face shield on the victim’s face. (921)
5.1.2	consumeren (to consume)	28 (2.0%)	Give a child something lukewarm with a lot of sugar to drink. (988)
5.1.3	verwijderen (to remove)	22 (1.6%)	Remove any (medicine) plasters from the victim. (957)
5.1.4	trekken (to pull)	25 (1.8%)	Carefully separate the eyelids with thumb and index finger. (935)
5.1.5	duwen (to push)	38 (2.7%)	Push the victim on his side. (905)
6.	onderzoeken (to investigate)	20 (1.4%)	Judge his breathing and start resuscitating. (914)
6.1	waarnemen (to observe)	32 (2.3%)	Moreover, the emergency officer on the phone will hear you. (954)
6.1.1	zien (to see)	54 (3.8%)	Judge his breathing by looking, listening and feeling for 10 secs. (911)
7.	contacteren (to contact)	18 (1.3%)	Otherwise, call the GP’s emergency number or call the GP centre. (982)
8.	schoonmaken (to clean)	15 (1.1%)	Rinse the victim’s eye for 15 minutes with lukewarm water. (972)

DISCUSSION & FUTURE WORK

- Improvement of existing tools to automatically analyse the Dutch MI dataset:
 - Alpino does not recognise all verbs + does not identify negations, conditions and warnings,
 - Cornetto provides hyperonyms for only two thirds of the MI lemma’s.
- Categorisation of the verbs in the PAT MI corpus to eventually automatically recognise actions in text and pictures.
- Comparison of matching textual and pictorial presentations using
 - more detailed analyses including aspect, modality, and adverbial specifications of manner
 - more fine-grained visual analysis identifying posture, gaze, and the positions of persons.
- Conducting user studies to test the effects of possible text-picture pairings.